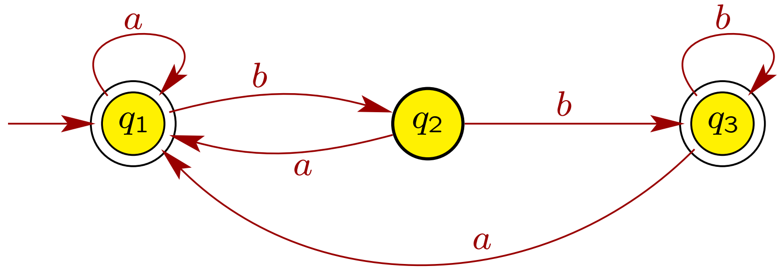
Homework 2

1. For the state diagram of the DFA M below, give its formal definition as a 5-tuple.



2. For each of the following languages over the alphabet $\Sigma = \{a, b\}$, give a DFA that recognizes the language. Give both a state diagram and 5-tuple specification for each DFA.

   (a) $A = \{\varepsilon, b, ab\}$.

   (b) For any string $w \in \Sigma^*$, let $n_a(w)$ denote the number of a's in w. For example, $n_a(abaaba) = 4$. Define the language

   $$B = \{ w \in \Sigma^* \mid n_a(w) \mod 3 = 1 \},$$

   i.e., $w \in B$ if and only if the number of a's in w is $3k + 1$ for some $k \geq 0$.

   (c) $C = \{ w \in \Sigma^* \mid w = saba \text{ for some string } s \in \Sigma^* \}$, i.e., C consists of strings that end in aba.

   (d) $D = \overline{C}$, where C is the language in the previous part; i.e., D consists of strings that do not end in aba.

   (e) $E = \{ w \in \Sigma^* \mid w \text{ begins with } b \text{ and ends with } a \}$.

   (f) For any string $w \in \Sigma^*$, let $n_b(w)$ denote the number of b's in w. Define the language $F = \{ w \in \Sigma^* \mid n_a(w) \geq 2, n_b(w) \leq 1 \}$.

   (g) $G = \{ w \in \Sigma^* \mid |w| \geq 2, \text{ second-to-last symbol of } w \text{ is } b \}$. If string $w = w_1 w_2 \cdots w_n$ where each $w_i \in \Sigma$, then the second-to-last symbol of w is w_{n-1}.

3. Show that, if M is a DFA that recognizes language B, swapping the accept and non-accept states in M yields a new DFA that recognizes $\overline{B}$, the complement of B. Conclude that the class of regular languages is closed under complementation.

4. We say that a DFA M for a language A is *minimal* if there does not exist another DFA M' for A such that M' has strictly fewer states than M. Suppose that $M = (Q, \Sigma, \delta, q_0, F)$ is a minimal DFA for A. Using M, we construct a DFA $\overline{M}$ for the complement $\overline{A}$ as $\overline{M} = (Q, \Sigma, \delta, q_0, Q - F)$. Prove that $\overline{M}$ is a minimal DFA for $\overline{A}$.

5. Give a formal proof that the class of regular languages is closed under intersection.

6. Let $\Sigma = \{ \texttt{a}, \texttt{b}, \ldots, \texttt{z}, \texttt{0}, \texttt{1}, \texttt{2}, \ldots, \texttt{9} \}$ be the alphabet consisting of lower-case Roman letters and Arabic numerals. Consider the language

$$L = \{ w \in \Sigma^* \mid w \text{ begins with a lower-case Roman letter} \}.$$

   (a) Give a DFA for L. For your DFA, give both a state diagram and 5-tuple for it.

   (b) Let J be the set of valid variable names in the Java programming language. Is $L \subseteq J$? Is $J \subseteq L$? Explain your answer.